



## PATENT ABSTRACTS OF JAPAN

(11) Publication number: **07261797 A**(43) Date of publication of application: **13 . 10 . 95**

(51) Int. Cl. **G10L 9/00**  
**G10L 7/06**  
**G10L 9/18**

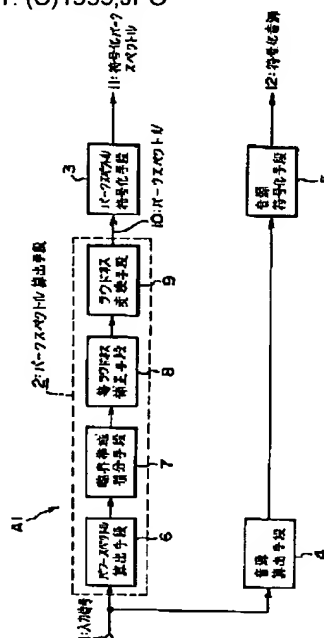
(21) Application number: **06049469**(71) Applicant: **MITSUBISHI ELECTRIC CORP**(22) Date of filing: **18 . 03 . 94**(72) Inventor: **TAZAKI HIROHISA****(54) SIGNAL ENCODING DEVICE AND SIGNAL DECODING DEVICE**

COPYRIGHT: (C)1995,JPO

**(57) Abstract:**

**PURPOSE:** To realize a signal encoding device and a signal decoding device which encodes and decodes a signal so as to be good for an auditory sense characteristic of a human with comparatively less operational quantity and to realize a signal encoding device which can suppress a noise component of a signal other than a voice signal with less operational quantity and memory quantity.

**CONSTITUTION:** A signal encoding device A1 is provided with a bark spectrum calculating means 2 which calculates a bark spectrum being a parameter based on an auditory sense model, a bark spectrum encoding means 3 which encodes a bark spectrum, a sound source calculating means 4, and a sound source encoding means 5. The bark spectrum calculating means 2 has a bark spectrum calculating means 6, a critical band integration means 7, an equal loudness correcting means 8 and a loudness conversion means 9. These means are formed by an engineering manner in order to obtain the same effect as the auditory sense model. When data is decoded, conversion is performed in the reverse direction.



(19) 日本国特許庁 (J P)

(12) 公開特許公報 (A)

(11)特許出願公開番号

特開平7-261797

(43)公開日 平成7年(1995)10月13日

(51)Int.Cl. <sup>6</sup>	識別記号	庁内整理番号	F I	技術表示箇所
G 1 0 L	9/00	J		
		G		
	7/06			
	9/18	E		

審査請求 未請求 請求項の数12 O L (全 17 頁)

(21)出願番号 特願平6-49469

(22)出願日 平成6年(1994)3月18日

(71)出願人 000006013

三菱電機株式会社

東京都千代田区丸の内二丁目2番3号

(72)発明者 田崎 裕久

神奈川県鎌倉市大船五丁目1番1号 三菱  
電機株式会社情報システム研究所内

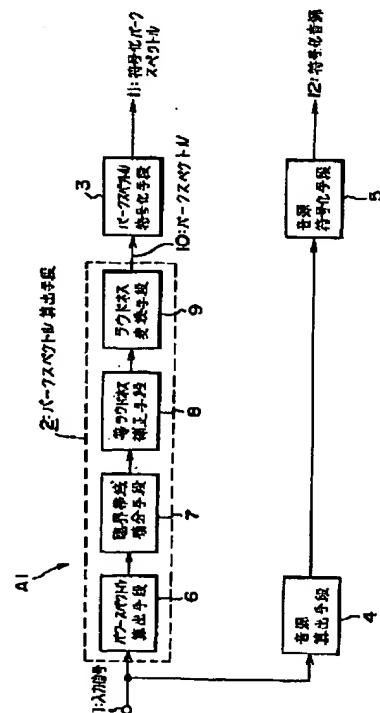
(74)代理人 弁理士 吉田 研二 (外2名)

(54) 【発明の名称】 信号符号化装置及び信号復号化装置

(57) 【要約】

【目的】 比較的少ない演算量で、人間の聴覚特性に対応の良い信号符号化及び信号復号化を実現することを目的としている。また、音声信号を符号化する場合、少ない演算量、メモリ量で音声信号以外の雑音の成分を抑圧できる信号符号化を実現することを目的としている。

【構成】 信号符号化装置A1に、聴覚モデルに基づくパラメータとしてのバークスペクトルを算出するバークスペクトル算出手段2とバークスペクトルを符号化するバークスペクトル符号化手段3と音源算出手段4と音源符号化手段5とを備える。バークスペクトル算出手段2は、パワースペクトル算出手段6と臨界帯域積分手段7と等ラウドネス補正手段8とラウドネス変換手段9の各手段を有し、これらの各手段は聴覚モデルと同様の効果を得るために工学的に生成されたものである。復号化に際しては、逆方向に変換を行う。



## 【特許請求の範囲】

【請求項 1】 入力信号に対して、聴覚モデルに基づくパラメータを算出し、聴覚モデルパラメータとして出力する聴覚モデルパラメータ算出手段と、前記聴覚モデルパラメータを符号化し、符号化聴覚モデルパラメータとして出力する聴覚モデルパラメータ符号化手段と、を備えたことを特徴とする信号符号化装置。

【請求項 2】 入力信号に対して、聴覚モデルに基づくパラメータを算出し、聴覚モデルパラメータとして出力する聴覚モデルパラメータ算出手段と、前記聴覚モデルパラメータを符号化し、符号化聴覚モデルパラメータとして出力する聴覚モデルパラメータ符号化手段と、符号化聴覚モデルパラメータを復号して、得られた復号聴覚モデルパラメータを出力する聴覚パラメータ復号化手段と、前記復号聴覚モデルパラメータを周波数スペクトル形状を表すパラメータに変換して、得られた周波数スペクトルパラメータを出力する変換手段と、複数の音源符号語を保持格納する音源符号帳と、前記復号聴覚モデルパラメータを用いて重み係数を算出し、前記音源符号帳内の各音源符号語に前記周波数スペクトルパラメータを乗じたものと入力音声の間の前記重み係数を用いた周波数領域での重み付き距離を計算し、この重み付き距離を最小にする音源符号語を選択して、その音源符号語を出力する音源符号選択手段と、を備えたことを特徴とする信号符号化装置。

【請求項 3】 聴覚モデルパラメータとしてバークスペクトルを用いたことを特徴とする請求項 1 又は 2 に記載の信号符号化装置。

【請求項 4】 入力信号に対して、有音区間であるか無音区間であるかの判定を行う有音無音判定手段と、無音区間の複数の前記聴覚モデルパラメータから、平均的な雑音の聴覚モデルパラメータを算出して、推定雑音パラメータとして出力する推定雑音パラメータ算出手段と、有音区間の前記聴覚モデルパラメータから前記推定雑音パラメータに相当する成分を取り除く雑音除去手段と、を備えたことを特徴とする請求項 1 から 3 のいずれかに記載の信号符号化装置。

【請求項 5】 聴覚モデルパラメータ算出手段が、入力信号のパワースペクトルを算出するパワースペクトル算出手段と、パワースペクトル算出手段で算出されたパワースペクトルに臨界帯域フィルタ関数を乗じて、励起パターンを算出する臨界帯域積分手段と、臨界帯域積分手段で算出された励起パターンに、周波数ごとの音の大きさと等感度との関係についての補正係数を乗じて補正励起パターンを算出する等ラウドネス補正手段

と、等ラウドネス補正手段で算出された補正励起パターンをパワー尺度からゾーン尺度へ変換してバークスペクトルを算出するラウドネス変換手段と、を有することを特徴とする請求項 3 に記載の信号符号化装置。

【請求項 6】 信号符号化装置が、さらに、入力信号に対して有音区間であるか無音区間であるかの判定を行う有音無音判定手段と、無音区間の複数の前記聴覚モデルパラメータから、平均的な雑音の聴覚モデルパラメータを算出して、推定雑音パラメータとして出力する推定雑音パラメータ算出手段と、を有し、聴覚モデルパラメータ算出手段が、入力信号のパワースペクトルを算出するパワースペクトル算出手段と、パワースペクトル算出手段で算出されたパワースペクトルに臨界帯域フィルタ関数を乗じて、励起パターンを算出する臨界帯域積分手段と、臨界帯域積分手段で算出された励起パターンに、周波数ごとの音の大きさと等感度との関係についての補正係数を乗じて補正励起パターンを算出する等ラウドネス補正手段と、有音区間の場合に、補正励起パターンから前記推定雑音パラメータに相当する成分を取り除いて、雑音を除去した補正励起パターンを算出する雑音除去手段と、雑音を除去した補正励起パターンをパワー尺度からゾーン尺度へ変換してバークスペクトルを算出するラウドネス変換手段と、を有することを特徴とする請求項 1 から 3 のいずれかに記載の信号符号化装置。

【請求項 7】 聴覚モデルに基づくパラメータを符号化した符号化聴覚モデルパラメータを復号して、得られた聴覚モデルパラメータを出力する聴覚パラメータ復号化手段と、前記聴覚モデルパラメータを周波数スペクトル形状を表すパラメータに変換して、得られた周波数スペクトルパラメータを出力する変換手段と、前記周波数スペクトルパラメータを用いて合成信号を生成する合成手段と、を備えたことを特徴とする信号復号化装置。

【請求項 8】 聴覚モデルパラメータとしてバークスペクトルを用いたことを特徴とする請求項 7 に記載の信号復号化装置。

【請求項 9】 周波数スペクトルパラメータとして周波数スペクトル振幅値を用いたことを特徴とする請求項 7 又は 8 に記載の信号復号化装置。

【請求項 10】 変換手段が、バークスペクトルをゾーン尺度からパワー尺度に変換して補正励起パターンを算出するラウドネス逆変換手段と、

上記補正励起ボタンに、周波数ごとの音の大きさと等感度との関係についての補正係数の逆数を乗じて励起ボタンを算出する等ラウドネス逆補正手段と、上記励起ボタンと臨界帯域フィルタ関数とから、パワースペクトルを算出するパワースペクトル変換手段と、上記パワースペクトルの各成分の平方根を算出して周波数スペクトル振幅値を算出する平方根手段と、を有することを特徴とする請求項8又は9に記載の信号復号化装置。

【請求項11】 聴覚モデルパラメータとしてバークスペクトル、周波数スペクトルパラメータとして周波数スペクトル振幅値を用い、前記変換手段では、周波数スペクトル振幅値を、バークスペクトルと同じ次数の周波数スペクトル振幅代表値による近似式で表現し、この近似式によりバークスペクトルと周波数スペクトル振幅代表値の間に成立する連立方程式を解くことで、バークスペクトルを周波数スペクトル振幅代表値に変換し、この周波数スペクトル振幅代表値と前記近似式を用いて周波数スペクトル振幅値を算出するようにしたことを特徴とする請求項2に記載の信号符号化装置。

【請求項12】 聴覚モデルパラメータとしてバークスペクトル、周波数スペクトルパラメータとして周波数スペクトル振幅値を用い、前記変換手段では、周波数スペクトル振幅値を、バークスペクトルと同じ次数の周波数スペクトル振幅代表値による近似式で表現し、この近似式によりバークスペクトルと周波数スペクトル振幅代表値の間に成立する連立方程式を解くことで、バークスペクトルを周波数スペクトル振幅代表値に変換し、この周波数スペクトル振幅代表値と前記近似式を用いて周波数スペクトル振幅値を算出するようにしたことを特徴とする請求項7に記載の信号復号化装置。

【発明の詳細な説明】

【0001】

【産業上の利用分野】本発明は、音声信号や音響信号などのデジタル信号を高効率に符号化する信号符号化装置、及び復号化を行う信号復号化装置に関するものである。

【0002】

【従来の技術】音声信号や音響信号を少ない情報量に圧縮する信号符号化においては、予め設定した歪を最小にするように符号の選択を行うことが一般的であり、その歪尺度としては、人間の聴覚特性と対応の良いものが望ましい。また、音声信号の符号化の場合には、音声信号以外の雑音が重畳していれば、雑音の成分を抑圧できる方式が望ましい。

【0003】人間の聴覚特性は、周波数軸が非線形であり、低域ほど分解能が高く、高域ほど分解能が低いことが知られている。その分解能は臨界帯域幅と呼ばれ、この臨界帯域幅に基づく周波数軸はバークスケールと呼ばれる。

【0004】また、人間の聴覚特性は、信号のパワーと線形比例しない、音の大きさに関する感覚量、ラウドネスを持っており、等しいラウドネスを与える信号パワーは周波数によって若干異なっていることが知られている。信号パワーが大きい場合には、信号のパワーの指数関数に周波数毎に若干異なる係数を乗じてラウドネスが近似算出される。

【0005】さらに、人間の聴覚特性の一つとして、妨害音がある場合、その他の信号を聞きとることのできる最小可聴値が上昇してしまう、マスキング効果も良く知られている。マスキング効果の大きさは、妨害音の周波数に近い周波数ほど大きく、周波数差がバークスケールでどれだけの幅であるかによって与えられる。

【0006】このような人間の聴覚特性や、そのモデル化の詳細については、Eberhard Zwicker著、山田由紀子訳「心理音響学」、161～174頁、1992年、西村書店、に記述されているので省略する。

【0007】聴覚特性と対応の良い歪尺度を用いた信号符号化装置として考えられたものとしては、例えば特開平4-55899号、特開平5-268098号、特開平5-15849号がある。

【0008】特開平4-55899号は、音声信号のスペクトルパラメータの符号化に聴覚特性と対応の良い歪尺度を導入したものである。まず、音声信号のスペクトル包絡を全極モデルで近似し、そのモデルのパラメータをスペクトルパラメータとして抽出する。そして、このスペクトルパラメータに対しメル変換などの非線形変換を行った後、2乗距離を歪尺度として符号化を行う。人間の聴覚特性の周波数軸の非線形性を、メル変換によって導入している。

【0009】特開平5-268098号は、短期予測と長期予測によって音声信号のスペクトル形状をほとんど取り去った残差信号の符号化に、バークスケールを導入したものである。残差信号を周波数領域に変換し、得られた全周波数成分を複数個ずつ結合して、バークスケールで等間隔におかれた結合振幅のみで表し、この結合振幅の符号化を行う。結合振幅を導入したことで周波数軸がバークスケールに疑似的に変換され、符号化時の歪、つまり結合振幅での歪が聴覚特性に対応が良くなる効果がある。

【0010】特開平5-158495号は、特性の異なる聴覚重み付けフィルタを持つ複数の音声符号化を行い、雑音感を最も低くする聴覚重み付けフィルタを選択するものである。雑音感の評価法の実施例として、入力音声信号と合成音信号の誤差を算出し、この誤差の入力音声信号に対するラウドネス、つまり雑音のラウドネスを求めることが記載されている。このラウドネスの算出には、臨界帯域幅、マスキング効果も導入されている。

【0011】さらに、聴覚特性と対応のよい歪尺度を開示した従来法としては、S. Wang, A. Sekey, A. Gersho 著

“Auditory Distortion Measure for Speech Coding”(Proc. ICASSP' 91, pp. 493-496, May 1991)がある。

【0012】S.Wangらの方法は、周波数スペクトル振幅に対して臨界帯域積分、等ラウドネス補正のためのプリエンファシス、ラウドネスへ変換するゾーン変換を行って得られる、バークスペクトルと呼ばれるパラメータを用いる。入力音声信号のバークスペクトルと、合成音声信号のバークスペクトルを算出し、この2つのバークスペクトルの単純な2乗誤差によって、入力音声信号と合成音声信号の間の歪を評価する方法である。なお、臨界帯域積分が上述した聴覚特性の周波数軸の非線形性とマスキング効果をモデル化するもので、プリエンファシスとゾーン変換が上述した聴覚特性のラウドネスに関する特性をモデル化するものである。

【0013】また、音声信号に重畳している雑音を抑圧する方式として考えられたものとしては、S.F.Boll著“Suppression of Acoustic Noise in Speech Using Spectral Subtraction”(IEEE Trans. on Acoustics, Speech, and Signal Processing, vol. ASSP-27, No. 2, pp. 113-120, April 1979)がある。

【0014】S.F.Bollの方法は、無音区間から雑音のスペクトル形状を推定し、これを全区間のスペクトルから減算することで、雑音成分の抑圧を行う方法であり、処理の概略は以下の通りである。

【0015】まず、入力信号を一定時間毎にハニング窓で切り出して、FFT(高速フーリエ変換)によって周波数スペクトルに変換する。次に、周波数スペクトルの各成分のパワーを計算してパワースペクトルを求める。無音区間と判定した区間において求められたパワースペクトルの平均をとることで、平均的な雑音のパワースペクトルを推定する。そして、この雑音のパワースペクトルに一定のゲインを乗じた後、これを全ての区間のパワースペクトルから減算する。この時、雑音の変動成分が雑音減算によって顕在化して、かえって雑音感が増加してしまう場合が多いので、減算によって非常に小さい値となった成分は、前後の切り出し区間の減算後の値との平滑化を行う。入力信号の周波数スペクトルと同じ位相スペクトルを持ち、前記平滑処理後のパワースペクトルと同じパワースペクトルを持つ周波数スペクトルを求めて、これを逆FFTして信号に戻す。最後に得られた信号を一定時間毎に接続していくことで復号信号とする。

【0016】

【発明が解決しようとする課題】上記の従来法には、以下に述べる課題がある。

【0017】すなわち、特開平4-55899号では、音声信号のスペクトル包絡を全極モデルで近似するが、この全極モデルは音声信号の生成機構モデルに基づくものである。全極モデルの最適パラメータ次数は、母音や子音、発声者によって異なるため、必ずしも良好な近似が行えるとは限らない。この点を改良する方法として、

最適パラメータ次数を推定して決定する方式の検討が行われたが、分析処理、合成処理の複雑化のためほとんど用いられていない。また、背景騒音等が重畳した音声信号に対しては、全極モデルの近似が合わなくなる課題もある。この従来法では、全極モデルに基づくパラメータに対して非線形変換を行って、聴覚特性に対応が良いように周波数軸を変換しているだけであるので、全極モデルの持つ上記課題は解消されない。また、ラウドネスや、マスキング効果等の他の聴覚特性に関するものが含まれていないので、聴覚特性との対応も十分ではない。また全極モデルは、音声信号以外の一般の音響信号に対しては適合しないので、この従来法を適用して、聴覚特性に対応の良い音響信号の符号化を行うことはできない。

【0018】また、メル変換の代わりに、全極モデルに基づくパラメータを一旦周波数スペクトル形状に変換し、これをバークスペクトルに変換することはできるので、全極モデルに基づくパラメータの符号化の符号化時の歪尺度にバークスペクトル歪を用いることが可能となるが、この変換が非常に多くの処理量を要するため、予め全符号に対して変換処理を行っておけるベクトル量子化を行う場合に限られる。また、前述した全極モデルに関する様々な課題が残されたままであるため、大幅な改善は期待できない。

【0019】また、特開平5-268098号では、残差信号の符号化にバークスケールを導入しているが、これも、聴覚特性の内、周波数軸の非線形性のみの導入であって、ラウドネスや、マスキング特性等の他の聴覚特性に関するものが含まれていないので、聴覚特性との対応は十分ではない。また、聴覚モデルは、耳に入力される信号そのものに適用して意味のあるものであり、この従来法のように適用対象を残差信号とした場合には、周波数軸の非線形変換以上の聴覚特性の導入は行えない。

【0020】また、特開平5-158495号では、聴覚重み付けフィルタの選択を行うための歪尺度として、雑音のラウドネスを用いているが、あくまでも聴覚重み付けフィルタの選択に用いるだけで、音声信号の符号化自体の歪尺度には用いていない。符号化の際の歪尺度は、聴覚重み付けフィルタ後の信号歪を用いているが、この聴覚重み付けフィルタは、全極モデルに基づいて、符号化によって発生する歪が聴感的に聞えにくいように周波数方向の歪重み付けを行うものであり、経験的に決定され、バークスケール、ラウドネス、マスキングといった聴覚特性は全く導入されていないものである。また、聴覚重み付けフィルタは、全極モデルのパラメータから導出されるので、音声信号以外の一般の音響信号に対しては適合しない。

【0021】この従来法の改良として、符号化の歪尺度に雑音のラウドネスを導入することが考えられるが、2のB乗種類(Bは符号化ビット数)存在する全ての符号

に対して合成音信号を生成し、その全てに対して雑音のラウドネスを算出しなければならないため、処理量が膨大で実現不可能である。

【0022】S.Wangらの方法では、聴覚モデルに基づくパラメータとしてバークスペクトルの算出を行うが、複数の合成音信号のバークスペクトル歪の評価を通じて、各種符号化方式の評価を行うことが目的であり、符号化の際の歪尺度としては用いることが検討されていない。仮に2のB乗種類（Bは符号化ビット数）存在する全ての符号に対して合成音信号を生成し、その全てに対してバークスペクトルを算出すれば、最もバークスペクトル歪が小さい符号を決定することができるが、処理量が膨大で実現不可能である。

【0023】S.F.Bollの方法では、雑音抑圧のために入力音声を一定時間毎にハニング窓で切り出して処理を行っているが、このハニング窓長と時間間隔は、FFTとのかねあいで2のべき乗の値となる。音声符号化装置でも入力音声を一定時間毎に切り出して処理を行うが、この時間間隔は必ずしも雑音処理の時間間隔と一致させられないため、雑音抑圧を完了した後に独立して音声符号化を行うこととなり、処理量、メモリ量を多く必要とし、信号のバックファイリングも複雑になる。時間間隔が一致した場合でも、少なくともFFTのポイント数（256、512、1024など）に比例した多くの演算処理とメモリが必要である。

【0024】また、実際には雑音減算によって雑音成分が低減されているにもかかわらず、変動分が顕在化してかえって雑音感が増すという聴覚的な課題を、単純なスペクトルの平滑化処理により改善を行っているのみであり、雑音の様態によっては改善が不十分である。

【0025】この発明は、かかる課題を解決するためになされたものであり、比較的少ない演算量で、人間の聴覚特性に対応の良い信号符号化及び信号復号化を実現することを目的としている。また、音声信号を符号化する場合、音声信号以外の雑音が重畳していれば、少ない演算量、メモリ量で雑音の成分を抑圧し、雑音の変動の影響の少ない聴覚的に良好な雑音抑圧ができる信号符号化を実現することを目的としている。

#### 【0026】

【課題を解決するための手段】この発明の請求項1に係わる信号符号化装置は、入力信号に対して、聴覚モデルに基づくパラメータを算出し、聴覚モデルパラメータとして出力する聴覚モデルパラメータ算出手段と、前記聴覚モデルパラメータを符号化し、符号化聴覚モデルパラメータとして出力する聴覚モデルパラメータ符号化手段とを備えたことを特徴とする。

【0027】また、請求項2に係る信号符号化装置は、入力信号に対して、聴覚モデルに基づくパラメータを算出し、聴覚モデルパラメータとして出力する聴覚モデルパラメータ算出手段と、前記聴覚モデルパラメータを符

号化し、符号化聴覚モデルパラメータとして出力する聴覚モデルパラメータ符号化手段と、符号化聴覚モデルパラメータを復号して、得られた復号聴覚モデルパラメータを出力する聴覚パラメータ復号化手段と、前記復号聴覚モデルパラメータを周波数スペクトル形状を表すパラメータに変換して、得られた周波数スペクトルパラメータを出力する変換手段と、複数の音源符号語を保持格納する音源符号帳と、前記復号聴覚モデルパラメータを用いて重み係数を算出し、前記音源符号帳内の各音源符号語に前記周波数スペクトルパラメータを乗じたものと入力音声の間の前記重み係数を用いた周波数領域での重み付き距離を計算し、この重み付き距離を最小にする音源符号語を選択して、その音源符号語を出力する音源符号選択手段と、を備えたことを特徴とする。

【0028】また、請求項3に係る信号符号化装置は、請求項1又は2の構成の信号符号化装置において、聴覚モデルパラメータとしてバークスペクトルを用いたことを特徴とするものである。

【0029】また、請求項4に係る信号符号化装置は、請求項1から3のいずれかの構成の信号符号化装置において、入力信号に対して、有音区間であるか無音区間であるかの判定を行う有音無音判定手段と、無音区間の複数の前記聴覚モデルパラメータから、平均的な雑音の聴覚モデルパラメータを算出して、推定雑音パラメータとして出力する推定雑音パラメータ算出手段と、有音区間の前記聴覚モデルパラメータから前記推定雑音パラメータに相当する成分を取り除く雑音除去手段と、を備えたことを特徴とするものである。

【0030】また、請求項5に係る信号符号化装置は、請求項3の構成の信号符号化装置において、聴覚モデルパラメータ算出手段が、入力信号のパワースペクトルを算出するパワースペクトル算出手段と、パワースペクトル算出手段で算出されたパワースペクトルに臨界帯域フィルタ関数を乗じて、励起パターンを算出する臨界帯域積分手段と、臨界帯域積分手段で算出された励起パターンに、周波数ごとの音の大きさと等感度との関係についての補正係数を乗じて補正励起パターンを算出する等ラウドネス補正手段と、等ラウドネス補正手段で算出された補正励起パターンをパワー尺度からゾーン尺度へ変換してバークスペクトルを算出するラウドネス変換手段と、を有することを特徴とするものである。

【0031】また、請求項6に係る信号符号化装置は、請求項1から3のいずれかの信号符号化装置において、信号符号化装置が、さらに、入力信号に対して有音区間であるか無音区間であるかの判定を行う有音無音判定手段と、無音区間の複数の前記聴覚モデルパラメータから、平均的な雑音の聴覚モデルパラメータを算出して、推定雑音パラメータとして出力する推定雑音パラメータ算出手段と、を有し、聴覚モデルパラメータ算出手段が、入力信号のパワースペクトルを算出するパワースペ

クトル算出手段と、パワースペクトル算出手段で算出されたパワースペクトルに臨界帯域フィルタ関数を乗じて、励起パターンを算出する臨界帯域積分手段と、臨界帯域積分手段で算出された励起パターンに、周波数ごとの音の大きさと等感度との関係についての補正係数を乗じて補正励起パターンを算出する等ラウドネス補正手段と、有音区間の場合に、補正励起パターンから前記推定雑音パラメータに相当する成分を取り除いて、雑音を除去した補正励起パターンを算出する雑音除去手段と、雑音を除去した補正励起パターンをパワー尺度からゾーン尺度へ変換してバークスペクトルを算出するラウドネス変換手段と、を有することを特徴とするものである。

【0032】また、請求項7の信号復号化装置は、聴覚モデルに基づくパラメータを符号化した符号化聴覚モデルパラメータを復号して、得られた聴覚モデルパラメータを出力する聴覚パラメータ復号化手段と、前記聴覚モデルパラメータを周波数スペクトル形状を表すパラメータに変換して、得られた周波数スペクトルパラメータを出力する変換手段と、前記周波数スペクトルパラメータを用いて合成信号を生成する合成手段と、を備えたことを特徴とするものである。

【0033】また、請求項8の信号復号化装置は、請求項7の信号復号化装置において、聴覚モデルパラメータとしてバークスペクトルを用いたことを特徴とするものである。

【0034】また、請求項9の信号復号化装置は、請求項7又は8に記載の信号復号化装置において、周波数スペクトルパラメータとして周波数スペクトル振幅値を用いたことを特徴とするものである。

【0035】また、請求項10の信号復号化装置は、請求項8又は9に記載の信号復号化装置において、変換手段が、バークスペクトルをゾーン尺度からパワー尺度に変換して補正励起パターンを算出するラウドネス逆変換手段と、上記補正励起パターンに、周波数ごとの音の大きさと等感度との関係についての補正係数の逆数を乗じて励起パターンを算出する等ラウドネス逆補正手段と、上記励起パターンと臨界帯域フィルタ関数とから、パワースペクトルを算出するパワースペクトル変換手段と、上記パワースペクトルの各成分の平方根を算出して周波数スペクトル振幅値を算出する平方根手段と、を有することを特徴とするものである。

【0036】また、請求項11の信号符号化装置は、請求項2に記載の信号符号化装置において、聴覚モデルパラメータとしてバークスペクトル、周波数スペクトルパラメータとして周波数スペクトル振幅値を用い、前記変換手段では、周波数スペクトル振幅値を、バークスペクトルと同じ次数の周波数スペクトル振幅代表値による近似式で表現し、この近似式によりバークスペクトルと周波数スペクトル振幅代表値の間に成立する連立方程式を解くことで、バークスペクトルを周波数スペクトル振幅

代表値に変換し、この周波数スペクトル振幅代表値と前記近似式を用いて周波数スペクトル振幅値を算出するようにしたことを特徴とするものである。

【0037】また、請求項12の信号復号化装置は、請求項7に記載の信号復号化装置において、聴覚モデルパラメータとしてバークスペクトル、周波数スペクトルパラメータとして周波数スペクトル振幅値を用い、前記変換手段では、周波数スペクトル振幅値を、バークスペクトルと同じ次数の周波数スペクトル振幅代表値による近似式で表現し、この近似式によりバークスペクトルと周波数スペクトル振幅代表値の間に成立する連立方程式を解くことで、バークスペクトルを周波数スペクトル振幅代表値に変換し、この周波数スペクトル振幅代表値と前記近似式を用いて周波数スペクトル振幅値を算出するようにしたことを特徴とするものである。

#### 【0038】

【作用】本発明における請求項1の信号符号化装置においては、聴覚モデルパラメータ算出手段が、バークスペクトル等の聴覚モデルに基づくパラメータを算出し、聴覚モデルパラメータ符号化手段が、このパラメータを直接符号化する。本発明に基づく信号符号化装置においては、聴覚モデルに基づくパラメータを直接符号化するようにしたので、聴覚特性との対応がよい信号符号化を行うことができる。また、請求項2の信号符号化装置においては、上記請求項1の場合と同様に、聴覚モデルパラメータ算出手段が聴覚モデルパラメータを出力し、聴覚モデルパラメータ符号化手段が、上記聴覚モデルパラメータを符号化して符号化聴覚モデルパラメータを出力する。さらに、聴覚パラメータ復号化手段が、上記符号化聴覚モデルパラメータを復号して復号聴覚モデルパラメータを出力し、変換手段が周波数スペクトルパラメータを出力する。音源符号選択手段は、上記復号モデルパラメータを用いて重み係数を算出し、上記音源符号帳内の各音源符号語に上記周波数スペクトルパラメータを乗じたものと入力信号間の重み付け距離を計算し、この重み付け距離を最小にする音源符号語を選択してその音源符号語を出力する。

【0039】本発明によれば、復号化後のパラメータを用いて算出した重み係数を用いて音源符号を探索するので、聴覚特性に対応のよい音源信号の符号化が可能となる。

【0040】また、請求項3の信号符号化装置においては、聴覚モデルパラメータとしてバークスペクトルを使用する。よって、パラメータ算出処理、符号化処理を少ない演算量で実現することができる。

【0041】また、請求項4の信号符号化装置においては、まず、有音無音判定手段が、入力信号に対して、有音区間か無音区間かを判定する。そして、推定雑音パラメータ算出手段が、無音区間の場合に、複数の聴覚モデルパラメータから平均的な雑音の聴覚モデルパラメータ

10

20

30

40

50

を算出して、推定雑音パラメータとして出力する。雑音除去手段は、有音区間の聴覚パラメータから推定雑音パラメータに相当する成分を取り除く。以上のようにして、雑音成分の抑圧を行って抑圧後の聴覚モデルパラメータを符号化する。

【0042】したがって、信号符号化とは非独立して雑音抑圧処理が実行できるようになり、また、雑音抑圧のための演算量とメモリ量を節約することができる。

【0043】また、請求項5の信号符号化装置においては、聴覚モデルパラメータ算出手段が、パワースペクトル算出手段と、臨界帯域積分手段と、等ラウドネス補正手段と、ラウドネス変換手段とを有し、まず、パワースペクトル算出手段が、入力信号のパワースペクトルを算出する。臨界帯域積分手段は、上記パワースペクトルに臨界帯域フィルタ関数を乗じて励起ボタンを算出する。等ラウドネス補正手段は、上記励起ボタンに、周波数ごとの音の大きさと等感度との関係についての補正係数を乗じて補正励起ボタンを算出する。次に、ラウドネス変換手段は、上記補正励起ボタンをパワー尺度からゾーン尺度へ変換してバークスペクトルを算出する。

【0044】本発明における信号符号化装置においては、臨界帯域積分手段によりマスキング効果が導入され、等ラウドネス補正手段により等ラウドネスの特性が導入される。また、ラウドネス変換手段によりゾーン尺度の特性が導入されるので、聴覚特性との対応がよい信号の符号化が可能となる。

【0045】また、請求項6の信号符号化装置においては、雑音除去手段が、等ラウドネス補正手段とラウドネス変換手段の間に設けられ、この雑音除去手段が、上記補正励起ボタンから推定雑音パラメータに相当する成分を取り除く。したがって、ラウドネス変換手段においてはパワー尺度からゾーン尺度へ変換する際に指数関数変換を行うことになるので、等ラウドネス補正手段が出力する励起ボタンから雑音を除去するようにしたことにより、演算処理を容易にすることができる。

【0046】また、請求項7の信号復号化装置においては、聴覚パラメータ復号化手段が、符号化聴覚モデルパラメータを復号して、得られた聴覚モデルパラメータを出力する。変換手段は周波数スペクトルパラメータを出力し、合成手段は、上記周波数スペクトルパラメータを用いて合成信号を生成する。本発明においては、符号化聴覚モデルパラメータを復号して、得られた周波数スペクトルパラメータを用いて合成信号を生成するので、聴覚特性との対応がよい信号復号化を行うことができる。

【0047】また、請求項8の信号復号化装置においては、聴覚モデルパラメータとしてバークスペクトルを用いたので、周波数スペクトルパラメータへの逆変換を少ない処理量で行うことができる。

【0048】また、請求項9の信号復号化装置においては、周波数スペクトルパラメータとして周波数スペク

トル振幅値を用いるので、様々な合成方法に適用が容易となる。

【0049】また、請求項10に記載の信号復号化装置においては、変換手段が、ラウドネス逆変換手段と、等ラウドネス逆変換手段と、パワースペクトル変換手段と、平方根手段とを有し、まず、ラウドネス逆変換手段が、バークスペクトルをゾーン尺度からパワー尺度へ変換して補正励起ボタンを算出する。次に、等ラウドネス逆補正手段が、上記補正励起ボタンに補正係数の逆数を乗じて励起ボタンを算出する。次に、パワースペクトル変換手段が、上記励起ボタンと臨界帯域フィルタ関数とからパワースペクトルを算出する。そして、平方根手段が、上記パワースペクトルの各成分の平方根を算出して周波数スペクトル振幅値を算出する。

【0050】本発明においては、ラウドネス逆変換手段によりゾーン尺度の特性が取り除かれ、等ラウドネス逆補正手段により等ラウドネスの特性が取り除かれ、かつ、パワースペクトル変換手段により、臨界帯域フィルタ関数の特性が取り除かれるので、聴覚特性との対応がよい信号の復号化が可能となる。

【0051】また、請求項11及び12に記載の信号符号化装置並びに信号復号化装置においては、バークスペクトルと同じ次数の周波数スペクトル振幅代表値による近似式で周波数スペクトル振幅値を表現することで、バークスペクトルから周波数スペクトル振幅値への近似変換を行うので、少ない処理量でバークスペクトルから周波数スペクトル振幅値への変換が実行できる。

【0052】

【実施例】

30 実施例1. 本発明の一実施例を図に基づいて説明する。

【0053】図1は本発明の一実施例である信号符号化装置A1の構成図である。図において、1は入力信号、2はバークスペクトル算出手段、3はバークスペクトル符号化手段、4は音源算出手段、5は音源符号化手段、6はパワースペクトル算出手段、7は臨界帯域積分手段、8は等ラウドネス補正手段、9はラウドネス変換手段、10はバークスペクトル、11は符号化バークスペクトル、12は符号化音源である。

40 【0054】ここで、バークスペクトル算出手段2は、パワースペクトル算出手段6と、上記パワースペクトル算出手段6に接続される臨界帯域積分手段7と、上記臨界帯域積分手段7に接続される等ラウドネス補正手段8と、上記等ラウドネス補正手段8に接続されるラウドネス変換手段9とを有し、バークスペクトル符号化手段3は上記ラウドネス変換手段9に接続されている。また、音源符号化手段5は、音源算出手段4に接続されている。

50 【0055】また、図2は本発明の一実施例である信号復号化装置Bの構成図である。図2において、11は符号化バークスペクトル、12は符号化音源、13はパー



クスベクトル復号化手段、14は変換手段、15は合成手段、16は音源復号化手段、17はラウドネス逆変換手段、18は等ラウドネス逆補正手段、19はパワースペクトル変換手段、20は平方根手段、21はバークスペクトル、22は周波数スペクトル振幅値、23は合成信号である。

【0056】ここで、変換手段14は、ラウドネス逆変換手段17と、ラウドネス逆変換手段17に接続される等ラウドネス逆補正手段18と、上記等ラウドネス逆補正手段18に接続されるパワースペクトル変換手段19と、上記パワースペクトル変換手段19に接続される平方根手段20とを有し、上記バークスペクトル復号化手段13は、上記ラウドネス逆変換手段17に接続されている。

【0057】ここで、信号符号化装置におけるバークスペクトル算出手段2は、人間の聴覚機構、すなわち、外耳、鼓膜、中耳、内耳、1次神経等の働きを工学的にモデル化したもので、聴覚モデルとも呼ばれる。聴覚モデル\*

$$D_i = \sum_j \{A_{ji} Y_i\}$$

ここで、臨界帯域フィルタ関数 $A_{ji}$ は、周波数 $i$ の信号が $j$ 番目の臨界帯域に与える刺激の強さを表す関数で、数式的なモデルの一例とその関数値のグラフがS. Wangらの文献に記載されている。マスキング効果は、この臨界帯域フィルタ関数 $A_{ji}$ に含まれて導入される。

【0061】次に、等ラウドネス補正手段8は、人間の聴覚が同じ音の強さと感じる振幅が周波数によって異なる、という特性を補正するために、励起パターン $D_j$ に補正係数 $H_j$ を乗じ、補正励起パターン $P_j$ を算出する。

【0062】次に、ラウドネス変換手段9は、補正励起パターン $P_j$ を、人間の感覚量としての音の大きさの尺度であるゾーン尺度に変換を行い、得られたパラメータをバークスペクトル10として出力する。そして、バークスペクトル符号化手段3は、バークスペクトル10を符号化して、得られた符号化バークスペクトル11を出力する。

【0063】なお、バークスペクトル符号化手段3における符号化には、スカラ量子化、ベクトル量子化、ベクトルスカラ量子化、多段ベクトル量子化、時間的に近接する複数のバークスペクトルを一括したマトリクス量子化などの方法を用いることが可能で、そこで用いる歪尺度としては、2乗距離や重み付き2乗距離などが良い。重み付き2乗距離における重み関数は、バークスペクトルの値の大きい次数に重みを強くする、時間的に前後との間のバークスペクトルの変化が大きい次数に重みを強くするなどである。

【0064】なお、上記構成においては、入力信号からバークスペクトルを算出するのに、パワースペクトル算出手段6と、臨界帯域積分手段7と、等ラウドネス補正手段8と、ラウドネス変換手段9とを使用しているが、

\*ルとしては、更に詳細なものも報告されているが、算出処理量が少ないことから、上記臨界帯域積分手段7と、等ラウドネス補正手段8と、ラウドネス変換手段9によって構成したものを用いる。

【0058】以下、図1と図2に示した本発明の一実施例の動作について説明する。

【0059】まず、例えば8KHzでサンプリングされたデジタル音声信号が、入力信号1としてバークスペクトル算出手段2内のパワースペクトル算出手段6に入力される。パワースペクトル算出手段6は、入力信号1に対してFFT（高速フーリエ変換）等のスペクトル変換処理を行い、得られた周波数スペクトル振幅値を2乗してパワースペクトル $Y_i$ を算出する。臨界帯域積分手段7は、式(1)に従って、パワースペクトル $Y_i$ に所定の臨界帯域フィルタ関数 $A_{ji}$ を乗じ、励起パターン $D_j$ を算出する。

【0060】

【数1】

(1)

この構成には限られず、例えば、臨界帯域積分手段7における臨界帯域積分関数に等ラウドネス補正手段8における補正係数を含めて構成することもでき、また、アナログ回路で構成することも可能である。また、符号化に際しては、ラウドネス変換手段9の出力を符号化しなくても、例えば、等ラウドネス補正手段8から出力される補正励起パターンを符号化したり、臨界積分手段7から出力される励起パターンを符号化してもよい。

【0065】一方、音源算出手段4は、まず、入力信号1が有声音であるか無声音であるかを判定し、その結果を有声音無声音判定結果とする。また、有声音と判定した場合には、ピッチ周波数を算出する。そして、この有声音無声音判定結果とピッチ周波数をまとめて、音源情報として出力する。音源符号化手段5は、この音源情報を符号化して、符号化音源12として出力する。

【0066】次に、信号復号化装置Bにおけるバークスペクトル復号化手段13は、前記符号化バークスペクトル11を復号化して、得られたバークスペクトル21を出力する。ここで、バークスペクトル復号化手段13における復号化は、前記バークスペクトル符号化手段3における符号化と対をなす方法で行う。つまり、バークスペクトル符号化手段3が所定の符号帳を用いたベクトル量子化を行う場合には、バークスペクトル復号化手段13においても同じ符号帳を用いた逆ベクトル量子化を行う、などとする。

【0067】変換手段14内のラウドネス逆変換手段17は、ラウドネス変換手段9の逆変換に相当し、ゾーン尺度からパワー尺度に戻し、補正励起パターン $P_j$ として出力する。等ラウドネス逆補正手段18は、等ラウドネス補正手段8の逆変換に相当し、補正励起パターン $P_j$ に

補正係数 $H_j$ の逆数を乗じることにより励起パタン $D_j$ を算出する。パワースペクトル変換手段19は、臨界帯域積分手段7の逆変換に相当し、励起パタン $D_j$ と帯域フィルタ関数 $A_{jk}$ から、後述する方法等によってパワースペクトル $Y_i$ を算出する。平方根手段20は、パワースペクトル $Y_i$ の各成分の平方根をとることで、周波数スペクトル振幅値22を算出する。

【0068】音源復号化手段16は、符号化音源12を復号化し、得られた音源情報を合成手段15に出力する。そして、合成手段15は、この音源情報と周波数スペクトル振幅値22を用いて、合成信号23の合成を行う。この合成の方法は、ハーモニックコーダの合成処理と同じものを用いることができる。同業者間では、一般的な方法であるので、ここでは詳細な説明は行わない。

【0069】なお、音源情報に有声音無声音判別結果とピッチ周波数を用いたが、この他に帯域別有声音無声音判別結果を加えて、多帯域励振(MBE)方式を用いて合成を行うなど、様々な変形が可能である。

【0070】音声信号や音響信号の場合、励起パタン $D_j$ の次数は15~24であり、パワースペクトル $Y_i$ の次数の方が大きい。このため、パワースペクトル変換手段19における変換処理は、一意に結果の得られるものではない。変換処理として、最も単純な方法は、図3に処理の流れを示すような、ニュートンラプソン法などの逐次求解法を用いるものである。

【0071】以下、図3に示した逐次求解法について説明する。

【0072】パワースペクトル変換手段14内には、臨界帯域積分手段7と同じものを備えさせる。そして、予め、臨界帯域フィルタ関数 $A_{jk}$ を用いて、パワースペクトル $Y_i$ の各成分に対する励起パタン $D_j$ の偏微分を算出しておく(ステップS1)。励起パタン $D_j$ が入力されたら(ステップS2)、まず仮のパワースペクトル $Y_i'$ を適切な初期値に設定する(ステップS3)。そして、この仮のパワースペクトル $Y_i'$ から、臨界帯域積分手段5と同じものを用いて仮の励起パタン $D_j'$ を算出し(ステップS4)、入力された励起パタン $D_j$ との誤差を算出する(ステップS5)。この誤差の2乗総和が所定の値 $\epsilon$ より小さい場合には、その時の仮のパワースペクトル $Y_i'$ をパワースペクトル $Y_i$ として出力する(ステップS6)。誤差の2乗総和が所定の値 $\epsilon$ 以上の場合には、この誤差と、予め算出しておいた偏微分を用いて、仮のパワースペクトル $Y_i'$ を更新する(ステップS7)。そして、ステップS4に戻る。

【0073】このように構成することにより、周波数軸の非線形性、感覚量であるラウドネス、マスキング効果といった聴覚特性を取り込んだ聴覚モデルに基づくパラメータを直接符号化・復号化することが可能となり、従来法に比べて、聴覚特性、つまり合成信号の主観品質との対応がよい信号符号化、信号復号化が実現される効果

がある。表現を変えれば、主観品質の劣化をできる限り抑えつつ、符号化情報量の削減が可能となる効果がある。

【0074】特にパークスペクトルは、少ない演算量で簡単に算出できること、その2乗距離や重み付き2乗距離などの簡単に計算できる距離尺度が主観歪と対応がよいこと、周波数スペクトル形状への逆変換が比較的少ない処理量で行えることから、このパークスペクトルを聴覚モデルに基づくパラメータとして用いることにより、パラメータ算出処理、符号化処理、変換処理が現実的な演算量で実現できる効果がある。

【0075】また、従来法を応用して聴覚モデルに基づくパラメータ上の歪を最小にしようとした場合のように、全ての符号に対する合成音の生成と聴覚モデルに基づくパラメータの算出を行うことがないので、現実的な演算量で信号符号化と信号復号化が実現できる効果がある。

【0076】さらに、従来法の課題であった、全極モデルによる近似、を排除することができているので、全極モデルのような最適次数の推定は不要であるし、背景雑音に強くなる効果がある。

【0077】また、周波数スペクトルパラメータとして、周波数スペクトル振幅値を用いるようにしたので、様々な合成処理が容易に適用できる効果がある。

【0078】実施例2。図4は本発明の他の一実施例である信号符号化装置A2の構成図である。図において新規な部分は、24のパークスペクトル復号化手段、25の変換手段、26の音源符号探索手段、27の音源符号帳がある。その他の部分は、図1と同等であり、説明を省略する。

【0079】以下、図4に示した本発明の一実施例の動作について説明する。

【0080】パークスペクトル復号化手段24は、図2において説明したパークスペクトル復号化手段13と同等のものであり、符号化パークスペクトル11を復号化して、得られたパークスペクトルを変換手段25に出力する。変換手段25は、図2において説明した変換手段14と同等のものであり、パークスペクトル復号化手段24が出力したパークスペクトルを周波数スペクトル振幅値に変換する。

【0081】音源符号探索手段26は、まず、入力信号1に対してFFT(高速フーリエ変換)等のスペクトル変換処理を行い、入力信号1の周波数スペクトル振幅値を得る。また、パワースペクトル $Y_i$ の各成分を微小変化させたときにパークスペクトルの2乗歪がどれだけ発生するかを算出し、重み係数 $G_i$ とする。次に、音源符号帳27内の全ての音源符号語を順番に読みだし、各々の音源符号語に対して、変換手段25が出力した周波数スペクトル振幅値を乗じ、これに適切なゲインを乗じたものと入力信号1の周波数スペクトル振幅値との間のG

、重み付き 2 乗距離を算出していく。そして、最も距離が小さかった音源符号語とゲインを選択し、その音源符号語に対応した音源符号語とゲインを符号化音源 12 として出力する。

【0082】重み係数  $G_i$  の算出は、例えば以下のようにして簡単に行える。まず、パワースペクトル  $Y_i$  の各成分に対する補正励起パターン  $P_i$  の偏微分を算出する。この偏微分は、臨界帯域フィルタ関数  $A_{jk}$  と等ラウドネス変換の係数から算出でき、固定的なものであるので、予め算出しておけば良い。次に、補正励起パターン  $D_j$  の各成分に微小摂動を与えたときにパークスペクトルの変化を算出し、その 2 乗総和を算出する。この値は、パークスペクトル復号化手段 24 が出力したパークスペクトルを変数とした簡単な式で計算できる。そして、算出されたパワースペクトル  $Y_i$  の各成分に対する補正励起パターン  $P_i$  の偏微分の行列と、補正励起パターン  $D_j$  の各成分に微小摂動を与えたときのパークスペクトルの変化の 2 乗総和を乗ずることで、所望の重み係数  $G_i$  が算出される。

【0083】ここでは説明を簡単にするために、入力信号 1 の周波数スペクトル振幅値を音源符号探索手段 26 内で算出する構成としたが、実際にはパークスペクトル算出手段 2 内のパワースペクトル算出手段 6 において既に算出しているので、これを記憶しておいて流用するほうが処理量が少なくて良い。

【0084】なお、この実施例によって得られた符号化結果は、図 2 の構成の信号復号化装置で復号化できる。ただし、音源復号化手段 16 と合成手段 15 の処理内容を変更する必要があるので、この部分のみ説明を行う。

【0085】音源復号化手段 16 は、符号化音源 12 を復号化し、得られた音源符号語とゲインを合成手段 15 に出力する。合成手段 15 は、音源符号語にゲインを乗じ、これにさらに周波数スペクトル振幅値 22 を乗じ、逆 FFT 等の逆フーリエ変換を行って、合成信号 23 を得る。

【0086】このように構成することにより、実施例 1 が持つ効果に加えて、聴覚特性に対応の良い音源信号の符号化・復号化が可能となる効果がある。また、聴覚モデルに基づくパラメータとしてパークスペクトルを用いた場合には、音源符号の探索に用いる重み係数が少ない処理量で算出できる効果がある。

【0087】実施例 3。図 5 は本発明の他の一実施例である信号符号化装置 A3 の構成図である。図において新規な部分は、有音無音判定手段 30、推定雑音パラメータ算出手段 31、雑音除去手段 32 がある。その他の部分は、図 1 と同等であり、説明を省略する。

【0088】以下、図 5 に示した本発明の一実施例の動作について説明する。

【0089】有音無音判定手段 30 は、入力信号 1 を分析して、音声区間であるか否か、すなわち有音か無音か

を判定し、結果を有音無音判定結果として出力する。推定雑音パラメータ算出手段 31 は、前記有音無音判定結果が無音の場合、等ラウドネス補正手段 8 が出力した補正励起パターンを用いて、内部に記憶してある推定雑音パターンを更新する。更新の方法としては、移動平均や、最近の無音の時の複数の補正励起パターンを記憶しておいて、その平均を算出しても良い。雑音除去手段 32 は、前記有音無音判定結果が有音の場合に、等ラウドネス補正手段 8 が出力した補正励起パターンから、前記推定雑音パラメータ算出手段 31 に記憶してある推定雑音パターンに所定のゲインを乗じたものを減算し、得られたものを新たに補正励起パターンとしてラウドネス変換手段 9 に出力する。

【0090】なお、上記雑音除去手段 32 では、有音の場合のみ減算処理を行っているが、無音の場合にも同じ減算処理を行うようにすることも可能であるし、無音の場合には等ラウドネス補正手段 8 が出力した補正励起パターンに 1.0 より小さいゲインを乗じたものを新たに補正励起パターンとしてラウドネス変換手段 9 に出力することも可能である。

【0091】このように構成することにより、実施例 1 が持つ効果に加えて、信号符号化と非独立に雑音抑圧処理が実行されるので、雑音抑圧のための演算量とメモリ量が節約でき、信号の複雑なバッファリングが不要となる効果がある。また、15 次程度のパークスペクトルの次数に比例した少ない演算量と少ないメモリ量で、S.F. Boll による従来例と同等の雑音抑圧効果が得られる。

【0092】さらに、従来は各周波数成分毎の減算処理であったので雑音の変動の影響が大きかったが、複数の周波数成分の積分を行って得られるパークスペクトルでは変動が平滑化されて小さくなっており、雑音の変動の影響が低減される効果がある。また、この平滑化は聴覚特性との対応が良く、従来の単純な平滑化処理に比べて良好な復号音質が得られる効果がある。

【0093】なお、雑音除去手段 32 は上記においては、等ラウドネス補正手段 8 とラウドネス変換手段 9 間に設けられているが、ラウドネス変換手段 9 の出力側に配置するようにしてもよい。

【0094】しかし、ラウドネス変換手段 9 においては、パワー尺度からゾーン尺度へ変換する際に指数関数変換を行うことになるので、ラウドネス変換手段 9 の出力側に配置すると、ラウドネス変換手段 9 における指数関数変換を考慮しなければならず、推定雑音パラメータ算出手段 31 で算出された雑音を単純に減算することができないという問題がある。よって、上記のように雑音除去手段 32 を等ラウドネス補正手段 8 とラウドネス変換手段 9 間に設けることにより、演算処理を簡単にすることができる。

【0095】実施例 4。上記実施例 3 では、実施例 1 の構成に有音無音判定手段 30、推定雑音パラメータ算出

手段31、雑音除去手段32を追加した構成となっているが、実施例2の構成に対して全く同様に有音無音判定手段30、推定雑音パラメータ算出手段31、雑音除去手段32を追加した構成も可能である。

【0096】このように構成することにより、実施例3が持つ効果に加えて、音源符号探索手段26で算出され、距離計算に用いられる重み係数が、雑音の比率の高い周波数において自動的に低くなり、合成音信号の了解性が改善される効果がある。

【0097】実施例5。上記実施例1乃至実施例4で、  
は、変換手段14及び変換手段25内のパワースペクトル\*

$$Y = RZ$$

ただし、

$$Y = [Y_1, Y_2, \dots, Y_N]^T$$

$$Z = [Z_1, Z_2, \dots, Z_N]^T$$

である。なお、Rとしては、RZ、つまりYが図6または図7に示すパターンとなるような行列を用いることができる。ここで、変数ベクトルZ<sub>j</sub>が周波数スペクトル振幅代表値に当たる。

※

$$D = AEY = AERZ$$

ただし、

$$D = [D_1, D_2, \dots, D_N]^T$$

である。

【0102】ここで、AERがM×M正方行列であるの☆

$$Y = R(AER)^{-1}D$$

従って、音源のパワースペクトルEが算出できれば、この式(4)を用いて、励起パターンDからパワースペクトルYへの変換が実行できる。

【0104】変換手段14内のパワースペクトル変換手段19に適用する場合には、音源復号化手段16が出力した音源情報を用いて音源のパワースペクトルを算出すれば良い。変換手段25内のパワースペクトル変換手段19に適用する場合には、まず一回前の時の音源を仮の音源として、そのパワースペクトルを算出してEとして用い、音源符号化探索手段26での探索を一回行って、その結果得られた音源のパワースペクトルを算出して再度パワースペクトル変換手段19での変換を行い、再度音源符号探索手段26での探索を行えば良い。仮の音源には、全極モデルによる残差信号や、入力信号1をケプストラム分析して、得られたケプストラムの20次程度以下の低次項を除いた後に、パワースペクトルに逆変換して用いても良い。

【0105】また、この近似求解法による変換を用いて算出されたパワースペクトルを、図3で説明した逐次求解法の初期値として、近似による誤差を低減することもできる。このように構成することにより、逐次求解法に比べて、少ない処理量でパークスペクトルから周波数スペクトル振幅値への変換が実行でき、信号符号化装置及び信号復号化装置の処理量が少なくなる効果がある。

【0106】実施例6。上記実施例1乃至実施例5にお

\*ル変換手段19においてニュートンラプソン法などの逐次求解法を用いて変換処理を行っていたが、これを以下に説明するような近似求解法を用いて行うこともできる。

【0098】近似求解法においては、最終的に算出するN次のパワースペクトルY<sub>i</sub>を、パークスペクトルと同じM次の変数ベクトルZ<sub>j</sub>と、予め固定的に与えた補間を表すM×N行列Rによって、式(2)のように近似して、求解を行う。

【0099】

【数2】

(2)

※【0100】この時、励起パターンD<sub>i</sub>は、音源のパワースペクトルを対角成分に持つN×N行列E、臨界帯域フィルタ関数A<sub>ji</sub>によって構成されるN×M行列Aを用いた式(3)のように表される。

【0101】

【数3】

(3)

☆で、逆行列が算出できる。式(2)と式(3)を変形して、下記の式(4)が導出される。

【0103】

【数4】

(4)

いて、パークスペクトル算出手段2内のパワースペクトル算出手段6と臨界帯域積分手段7を、臨界帯域フィルタの特性を模擬したバンドパスフィルタ群とパワーを積分する手段で構成することもできる。つまり、パラメータを抽出して符号化を行う周期は20msec(この周期をフレームと呼ぶ)であり、このフレーム内においては入力信号のスペクトルは定常であると考え、フレーム内のバンドパスフィルタの出力を積分していくのである。また、パワーを積分する手段はローパスフィルタで実現することもできる。また、等ラウドネス補正手段8をも含めた特性を与えても良い。

【0107】このように構成することにより、フィルタの次数が少なく、パークスペクトルの算出周期が短い場合には、処理量が削減される効果がある。

【0108】実施例7。上記実施例1乃至実施例6において、パークスペクトル符号化手段3が、時間的に近接する複数のパークスペクトルを記憶しておき、セグメント量子化を行う構成も可能である。セグメント量子化を行う場合には、セグメント境界の決定がその符号化特性に大きな影響を与えるが、パークスペクトルの時間的変化速度が極大または極小の部分を境界としたり、これを初期値としてパークスペクトルの符号化歪が最小になるように境界を決定したりすれば良い。

【0109】このように構成することにより、上記実施例1乃至実施例6の効果に加えて、セグメント境界の決

定が聴覚的に歪が小さくなるように決定できる効果がある。実施例 8、上記実施例 1 乃至請求項実施例 7 において、臨界帯域積分手段 7 における臨界帯域フィルタ関数、等ラウドネス補正手段 8 内の補正係数、ラウドネス変換手段 9 におけるパワー尺度からゾーン尺度への変換特性を複数組用意しておき、どの組を用いるかを使用者が選択できるように構成することもできる。複数組の一例としては、1 つは正常な聴覚特性を模擬する変換特性、臨界帯域フィルタ関数、及び補正係数、もう一つは若干劣化が起った老人の聴覚特性を模擬する変換特性、臨界帯域フィルタ関数、及び補正係数とする。この他、特定の難聴者に対して、その聴覚特性を測定して、その聴覚特性を模擬する変換特性、臨界帯域フィルタ関数、及び補正係数を含めて用いても良い。なお、どの組を用いたかを交換手段 14 及び交換手段 25 内のラウドネス逆変換手段 17、等ラウドネス逆補正手段 18、パワースペクトル変換手段 19 に伝達し、これらにおいて用いる変換特性、臨界帯域フィルタ関数、及び補正係数も連動させる。

【0110】このように構成することにより、老人等の聴覚特性が異なる聴取者に対しても、実施例 1 乃至実施例 7 と同様の効果が得られ、従来法に比べて、聴覚特性、つまり合成信号の主観品質との対応がよい信号符号化、信号復号化が実現される効果がある。

【0111】実施例 9、上記実施例 1 乃至実施例 8 の交換手段 14 において、ラウドネス逆変換手段 17 のゾーン尺度からパワー尺度への変換特性、等ラウドネス逆補正手段 18 の臨界帯域フィルタ関数、パワースペクトル変換手段 19 の補正係数を複数組用意しておき、どの組を用いるかを使用者が選択できる構成とすることもできる。複数組の一例としては、1 つは正常な聴覚特性を模擬する変換特性、臨界帯域フィルタ関数、及び補正係数、もう一つは若干劣化が起こった老人の聴覚特性を補償する変換特性、臨界帯域フィルタ関数、及び補正係数とする。この他、特性の難聴者に対して、その聴覚特性を測定して、その聴覚特性を補償する変換特性、臨界帯域フィルタ関数、及び補正係数を含めて用いても良い。

【0112】このように構成することにより、老人等の聴覚特性が異なる聴取者の聞き取り易い信号復号化が実現される効果がある。

#### 【0113】

【発明の効果】以上説明したように請求項 1 の発明は、聴覚モデルに基づくパラメータを算出し、このパラメータを直接符号化するようにしたので、従来法に比べて、聴覚特性との対応がよい信号符号化が実現される効果がある。表現を変えれば、主観品質の劣化をできる限り抑えつつ、符号化情報量の削減が可能となる効果がある。

【0114】また、従来法を応用して聴覚モデルに基づくパラメータ上の歪を最小にしようとした場合のよう

に、全ての符号に対する合成音の生成と聴覚モデルに基

づくパラメータの算出を行うことがないので、現実的な演算量で信号符号化が実現できる効果がある。

【0115】さらに、従来法の課題であった、全極モデルによる近似、を排除することができているので、全極モデルのような最適次数の推定は不要であるし、背景雑音に強くなる効果がある。

【0116】請求項 2 の発明は、聴覚モデルに基づくパラメータを算出し、このパラメータを直接符号化・復号化するとともに、復号化後のパラメータを用いて算出した重み係数を用いて音源符号の探索を行うようにしたので、請求項 1 の発明が持つ効果に加えて、聴覚特性に対応の良い音源信号の符号化が可能となる効果がある。

【0117】請求項 3 の発明は、請求項 1 乃至請求項 2 の信号符号化装置における聴覚モデルに基づくパラメータとして、パワースペクトルを用いるようにしたので、請求項 1 乃至請求項 2 の発明が持つ効果に加えて、パラメータ算出処理、符号化処理の少ない演算量で実現できる効果がある。

【0118】また、請求項 2 の信号符号化装置における請求項 3 の発明は、距離計算に用いられる重み係数が少ない処理量で算出できる効果もある。

【0119】請求項 4 の発明は、無音区間の聴覚モデルパラメータから雑音の平均的な聴覚モデルパラメータを推定し、これを有音区間の聴覚モデルパラメータから除去することで、雑音成分の抑圧を行い、抑圧後の聴覚モデルパラメータを符号化するようにしたので、請求項 1 乃至請求項 3 の発明が持つ効果に加えて、信号符号化と非独立に雑音抑圧処理が実行されるようになり、雑音抑圧のための演算量とメモリ量が節約でき、信号の複雑なバッファリングが不要となる効果がある。また、聴覚モデルパラメータとしてパワースペクトルを用いた場合には、15 次程度のパワースペクトルの次数に比例した少ない演算量と少ないメモリ量で、従来例と同等の雑音抑圧効果が得られる効果がある。

【0120】さらに、従来は各周波数成分毎の減算処理であったので雑音の変動の影響が大きかったが、聴覚モデルパラメータでは変動が周波数方向に平滑化されて小さくなっており、雑音の変動の影響が低減される効果がある。また、この平滑化は聴覚特性との対応が良く、従来の単純な平滑化処理に比べて良好な復号音質が得られる効果がある。

【0121】また、請求項 2 の信号符号化装置における請求項 4 の発明は、距離計算に用いられる重み係数が、雑音の比率の高い周波数において自動的に低くなり、合成音信号の了解性が改善される効果もある。

【0122】また、請求項 5 の発明においては、臨界帯域積分手段によりマスキング効果が導入され、等ラウドネス補正手段により等ラウドネスの特性が導入され、また、ラウドネス変換手段によりゾーン尺度の特性が導入されるので、聴覚特性との対応がよい信号の符号化が可

能となる。

【0123】請求項6に記載の発明においては、等ラウドネス補正手段が出力する励起パターンから雑音を除去するようにしたことにより、演算処理を容易にすることができる。

【0124】請求項7の発明は、聴覚モデルパラメータを周波数スペクトルパラメータに変換し、得られた周波数スペクトルパラメータを用いて合成音信号の生成を行うようにしたので、聴覚特性との対応がよい信号復号化が実現される効果がある。

【0125】請求項8の発明は、請求項7の信号復号化装置における聴覚モデルパラメータとして、バークスペクトルを用いるようにしたので、請求項7の発明が持つ効果に加えて、周波数スペクトルパラメータへの逆変換が比較的少ない処理量で行え、変換処理が現実的な演算量で実現できる効果がある。

【0126】請求項9の発明は、請求項7乃至請求項8の信号復号化装置における周波数スペクトルパラメータとして、周波数スペクトル振幅値を用いるようにしたので、請求項5乃至請求項6の発明が持つ効果に加えて、

様々な合成方式に適用が容易となる効果がある。

【0127】請求項10の発明においては、ラウドネス逆変換手段によりゾーン尺度の特性が取り除かれ、等ラウドネス逆補正手段により等ラウドネスの特性が取り除かれ、かつ、パワースペクトル変換手段により、臨界帯域フィルタ関数の特性が取り除かれるので、聴覚特性との対応がよい信号の復号化が可能となる。

【0128】請求項11と12の発明は、バークスペクトルと同じ次数の周波数スペクトル振幅代表値による近似式で周波数スペクトル振幅値を表現することで、バークスペクトルから周波数スペクトル振幅値への近似変換を行うようにしたので、少ない処理量でバークスペクトルから周波数スペクトル振幅値への変換が実行でき、信号符号化装置及び信号復号化装置の処理量が少なくなる効果がある。

#### 【図面の簡単な説明】

【図1】この発明の実施例1の信号符号化装置を示す構成図である。

【図2】この発明の実施例1の信号復号化装置を示す構成図である。

\*【図3】この発明の実施例1のパワースペクトル変換手段19における逐次求解法を説明するフローチャートである。

【図4】この発明の実施例2の信号符号化装置を示す構成図である。

【図5】この発明の実施例3の信号符号化装置を示す構成図である。

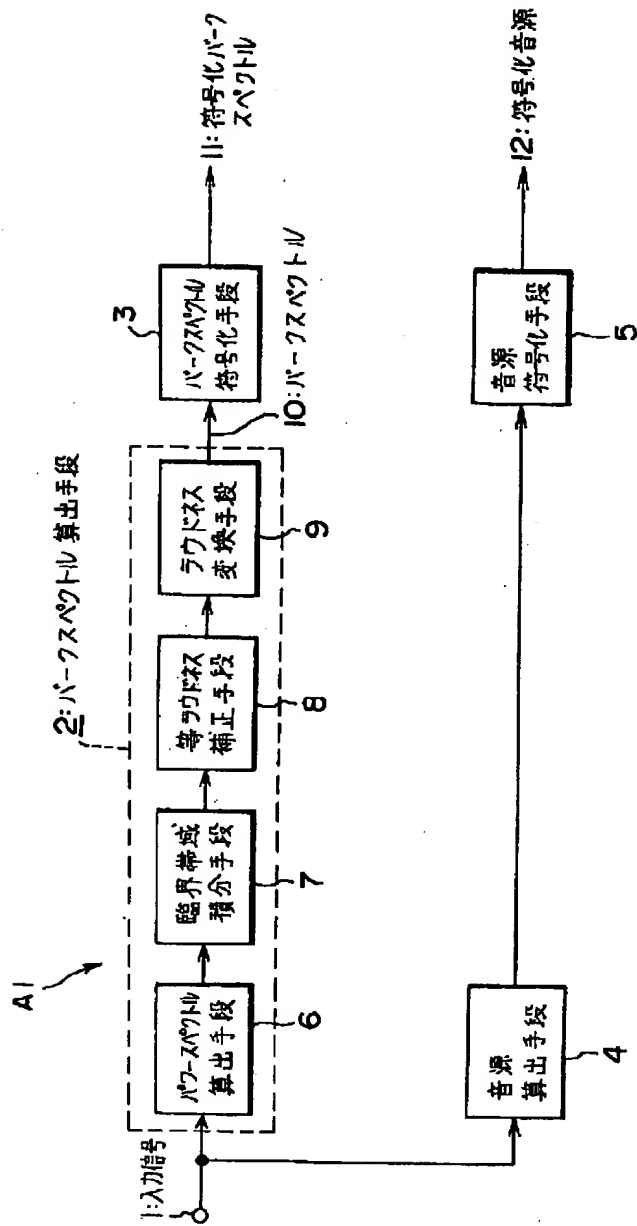
【図6】この発明の実施例5における補間を表す行列の一例を説明する概念図である。

10 【図7】この発明の実施例5における補間を表す行列の一例を説明する概念図である。

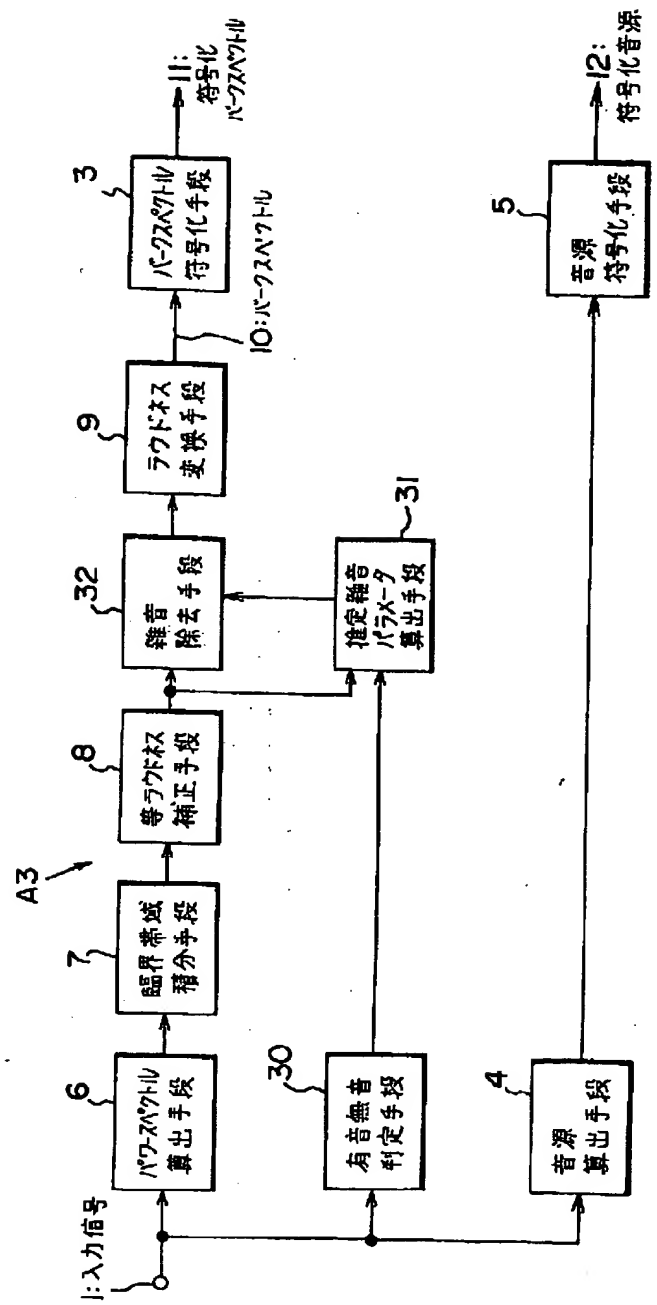
#### 【符号の説明】

- 1 入力信号
- 2 バークスペクトル算出手段
- 3 バークスペクトル符号化手段
- 4 音源算出手段
- 5 音源符号化手段
- 6 パワースペクトル算出手段
- 7 臨界帯域積分手段
- 20 8 等ラウドネス補正手段
- 9 ラウドネス変換手段
- 10、21 バークスペクトル
- 11 符号化バークスペクトル
- 12 符号化音源
- 13、24 バークスペクトル復号化手段
- 14、25 変換手段
- 15 合成手段
- 16 音源復号化手段
- 17 ラウドネス逆変換手段
- 30 18 等ラウドネス逆補正手段
- 19 パワースペクトル変換手段
- 20 平方根手段
- 22 周波数スペクトル振幅値
- 23 合成信号
- 26 音源符号探索手段
- 27 音源符号帳
- 30 有音無音判定手段
- 31 推定雑音パラメータ算出手段
- 32 雑音除去手段

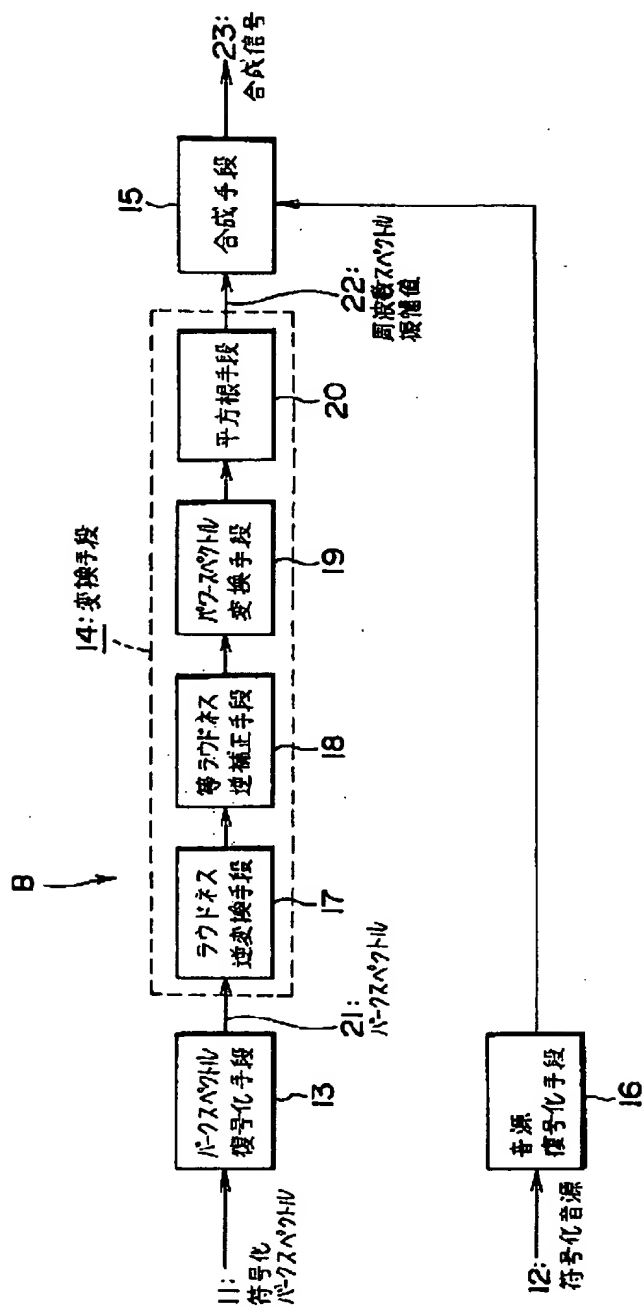
【図 1】



【図 5】

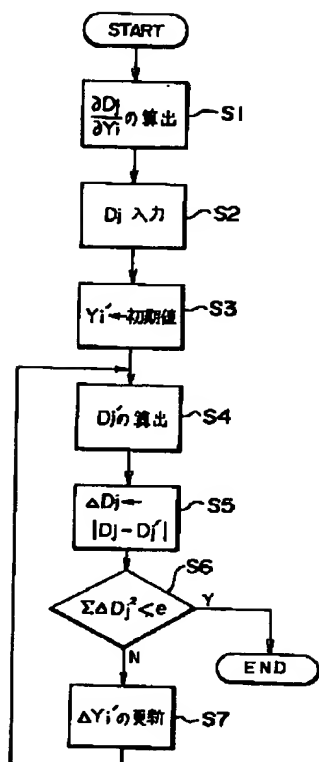


【図 2】

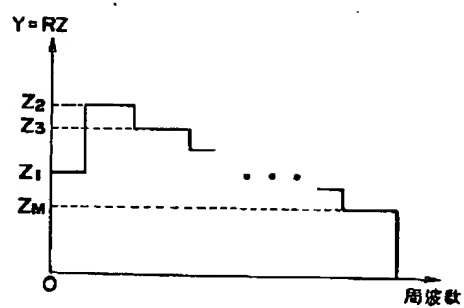




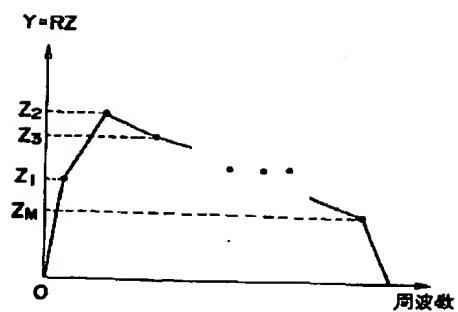
【図3】



【図6】



【図7】



【図 4】

